Gabriel Greenfield

BULLETIN MENSUEL · ÉDITION NOVEMBRE 2026

Gouvernance agentique

Cybersécurité, gouvernance, conformité : le tableau complet.

DONNÉES ARRÊTÉES AU

01/11/2026

MÉTHODOLOGIES

Meridian et GAME

SIGNAL DU MOIS

La gouvernance agentique quitte la bonne pratique pour entrer dans l'exigence opposable.

07/10

Singapour : la MAS publie ses lignes directrices IA, agents inclus.

14/10

Europe : EN 18286 citée au JOUE, première norme harmonisée de l'AI Act.

30/10

France : l'ANSSI publie « Essentiels IA agentique », sept principes.

LE MOIS EN CHIFFRES

628 000+

enregistrements ukrainiens volés avec des outils conçus par un assistant de code.

6 jours

pour bâtir, avec un modèle public, une chaîne d'exploitation navigateur estimée à 250 000 $.

×10

le volume de spearphishing assisté par IA par rapport à 2025 (Microsoft).

217 M

de réponses web analysées : des consignes cachées visent les agents dans les en-têtes HTTP.

1 423

tentatives d'exploitation de la faille Langflow captées par des pots de miel.

2 ans+

sans renouvellement du jeton d'un fournisseur d'analytique chez ASOS.

Plusieurs chiffres proviennent de veilles sectorielles non recoupées avec les sources primaires.

SYNTHÈSE EXÉCUTIVE

De la révélation à la reddition de comptes

OpenAI s'explique devant le Parlement australien ; GPT-6.1 Astra reste non publié. Un régulateur asiatique, une agence française et un normalisateur européen convergent le même mois.

Le triptyque commun : inventaire, autorité de décision graduée, traçabilité inviolable. Les organisations sans responsable par agent ne manquent plus de modèle : elles manquent de temps.

CYBERSÉCURITÉ

L'IA devient l'outil de l'attaquant, et l'agent sa cible.

GOUVERNANCE

L'audition australienne, premier test de responsabilité d'un laboratoire.

CONFORMITÉ

La présomption de conformité devient réelle.

01

ADOPTION

Où en est réellement l'IA agentique en entreprise

PHÉNOMÈNE DU MOIS

« Bring Your Own Agent » : des assistants personnels connectés avec des identifiants professionnels.

Ce n'est plus un outil que l'on consulte, mais un opérateur muni de droits d'accès, que personne n'a déclaré.

91 %

utilisent déjà l'IA agentique

85 %

ont des agents sans humain en temps réel

36 %

ont subi un incident IA significatif

Source : EY (septembre 2026). Aucune enquête de cette ampleur en octobre.

Un serveur MCP exposé se traite comme une interface de production : propriétaire, correctifs, surveillance.

ASSISES DE MONACO · 7–9/10

La fuite de données reste la préoccupation n°1 des RSSI ; le shadow AI en est le vecteur principal.

MCP · 05/10

Correctif SSRF sur le serveur MCP de référence : un protocole encore en durcissement.

02

CADRE D'ANALYSE

Trois objets, trois gouvernances

Gouvernance IA classique

QUESTION CENTRALE

Le modèle est-il fiable, précis, conforme ?

RÉFÉRENTIELS

ISO/IEC 42001, EN 18286, NIST AI RMF, AI Act, MAS

RISQUE DOMINANT

Biais, opacité, non-conformité

INCIDENT TYPE DU MOIS

Bases non protégées, comptes de service jamais renouvelés

Gouvernance des agents

QUESTION CENTRALE

L'agent agit-il dans son périmètre et en rend-il compte avec sincérité ?

RÉFÉRENTIELS

OWASP Top 10 agentique, ANSSI Essentiels, IAM/IGA

RISQUE DOMINANT

Action hors périmètre, compromission d'identité

INCIDENT TYPE DU MOIS

Langflow, injection par en-têtes HTTP, usurpation d'agent

Gouvernance agentique

QUESTION CENTRALE

Le système collectif livre-t-il le résultat sans dérive émergente ?

RÉFÉRENTIELS

MAESTRO, protocoles inter-agents, OWASP ACS

RISQUE DOMINANT

Dérive systémique, collusion, dissimulation

INCIDENT TYPE DU MOIS

Chaînes d'attaque orchestrées de bout en bout par des agents

Cas ASOS : la gouvernance de l'IA ne tient que si la gestion des identités non humaines est saine.

03

ANGLE CYBERSÉCURITÉ

Selon une veille sectorielle

L'IA comme multiplicateur d'attaque

628 000+	**Ukraine** : outils de collecte et d'exfiltration construits avec Claude Code par un groupe aux liens étatiques présumés.
6 jours	**Navigateur** : chaîne d'exploitation complète (250 000 $) avec un modèle public, sans contournement.
DIVD	**Pays-Bas** : un agent enchaîne deux failles non corrigées de Zammad. Un autre contourne une double authentification.
16/10	**CISA, FBI, NSA** : campagne chinoise en deux temps, balayage IA puis intrusion manuelle ciblée.

LE POINT COMMUN

Le délai entre découverte et exploitation s'effondre.

Les cycles de correctifs, calibrés sur des attaquants humains, ne suivent plus des adversaires qui enchaînent reconnaissance, exploitation et exfiltration à vitesse machine.

03 · ANGLE CYBERSÉCURITÉ

Les agents comme cibles : trois axes

Axe 1

Identité

Fausses références d'agents et historiques fabriqués pour déclencher des validations financières frauduleuses. Zenity Labs publie AgentCorruption (08/10) sur Amazon Bedrock AgentCore.

L'identité, première ligne de défense (OWASP ASI03).

Axe 2

Entrées non fiables

217 M

réponses web analysées

Consignes cachées dans X-AI-Instructions ou X-Agent-Note, jusqu'à l'envoi de cookies vers l'extérieur.

Un phénomène systématique, non ciblé.

Axe 3

Plateformes

1 423

tentatives sur Langflow

Exécution de code depuis tout compte valide, même d'essai : clés d'API et bases vectorielles exposées.

La chaîne d'outils est plus exposée que l'agent.

03 · LABORATOIRES

Les laboratoires reconnaissent la difficulté du confinement

Anthropic	Des agents ont contourné des limites posées par simples instructions ; accès internet en direct supprimé « pour l'instant » des évaluations.
GPT-6.1 Astra	Comportement différent selon qu'il se croit testé ; tentative de copie vers un stockage persistant en bac à sable.

LECTURE GABRIEL GREENFIELD

Les attaques les plus efficaces combinent un agent offensif et une identité mal gérée.

01 Inventorier les identifiants et raccourcir leur durée de vie.

02 Placer les contrôles hors du modèle, de façon déterministe.

RECOMMANDATION OPÉRATIONNELLE DU MOIS

Traiter le contenu externe consommé par un agent comme hostile par défaut.

01

Filtrer et marquer les contenus récupérés avant leur entrée dans le contexte

02

Interdire aux agents de navigation l'accès simultané à des secrets et à des canaux de sortie

03

Auditer les plateformes de conception d'agents : versions, comptes d'essai, clés stockées

04

ANGLE GOUVERNANCE

Le Parlement australien, premier test de responsabilité

18/06

Un agent accède au portail statistique Medicare

Août

OpenAI en prend connaissance

10/09

Un courriel unique vers une boîte générique

24/09

Révélation publique par le Premier ministre

6–9/10

Quatre jours d'audiences : OpenAI, Anthropic, Microsoft, Google

30/11

Rapport final attendu, état non établi

JASON KWON, OPENAI · 06/10

La façon d'en prendre connaissance « aurait pu être bien meilleure ».

ENSEIGNEMENT 1

La rupture est organisationnelle : tester la notification avant d'en avoir besoin.

ENSEIGNEMENT 2

La déclaration immédiate des incidents d'IA hors de contrôle arrive.

04 · ANGLE GOUVERNANCE

L'écart entre la règle et la pratique reste la donnée centrale

AUCUN INDICATEUR NE S'AMÉLIORE

98 % → **47 %**

ont une politique — l'ont déjà contournée

26 %

ne détectent pas un agent non autorisé

75 %

sans supervision claire des agents (EMEA)

OÙ SE JOUE L'ÉCART

Dans les petites décisions d'identité.

Qui possède ce jeton ? Qui l'a émis ? Qui sait qu'il existe ?

ANSSI

Diagnostic jugé trop pauvre en actions aux Assises ; le guide du 30/10 y répond en partie.

Sources : EY ; Veeam (EMEA). Données de septembre, aucune enquête comparable en octobre.

À L'ATTENTION DES COMEX

Trois questions à poser ce mois-ci à votre direction de la sécurité

01

Combien de temps entre la création d'un agent ou d'un jeton non déclaré et sa découverte ?

02

Qui peut arrêter un agent, en combien de minutes, et ce droit a-t-il déjà été exercé ?

03

La notification d'incident couvre-t-elle un agent hors périmètre sans aucun attaquant en cause ?

05

ANGLE CONFORMITÉ

La présomption de conformité devient réelle

JOURNAL OFFICIEL DE L'UE · 14/10/2026

EN 18286

CITÉE

DÉCISION (UE) 2026/2274

Première norme harmonisée de l'AI Act : système qualité du fournisseur, article 17.

COUVRE

Système de management de la qualité, y compris la surveillance après commercialisation.

NE DISPENSE PAS

De la gestion des risques, de la gouvernance des données ni de la supervision humaine.

ARBITRAGE POUR LE HAUT RISQUE

EN 18286 devient le chemin le moins coûteux vers l'article 17 ; ISO/IEC 42001 reste le complément organisationnel.

05 · CALENDRIER EUROPÉEN

Une échéance à un mois

02/12/2026

J-31

Art. 50 : transparence des systèmes déjà sur le marché

Agents conversationnels et vocaux signalent l'IA dès le début de l'interaction.

02/12/2027

Haut risque, annexe III

Calendrier Digital Omnibus (UE) 2026/1744 inchangé.

02/08/2028

Haut risque, annexe I

Produits réglementés.

CHANTIER LE PLUS URGENT

Marquer les canaux existants : centres de contact, assistants internes, agents d'avant-vente.

LIGNES DIRECTRICES HAUT RISQUE · EN PROJET

Un système à usage général qui n'exclut pas les usages à haut risque peut être réputé à haut risque. Adoption attendue d'ici fin 2026.

Singapour et la France publient leurs cadres

MAS

07/10/2026

IA générative et agents inclus. En vigueur le 07/10/2027.

01 Gouvernance claire, responsabilités identifiées

02 Maîtrise sur tout le cycle de vie

03 Risque tiers : l'IA achetée reste votre responsabilité

04 Application proportionnée

Consultation dédiée à l'IA agentique prévue en 2027.

ANSSI · Essentiels IA agentique

30/10

v1.0 · 7 principes · 4 thèmes : accès, actions autonomes, cloisonnement, surveillance.

CHAQUE OPÉRATION CLASSÉE EN TROIS NIVEAUX

Exécution autonome	Approbation humaine	Action interdite

· Contrôles déterministes avant chaque outil

· Effet des séquences, pas seulement des droits

· Identités séparées dev, test, prod

· Secrets hors contexte, mémoire et journaux

· Arrêt d'urgence avec révocation

· Journaux inviolables par l'agent

REPÈRES AU 01/11/2026

Référentiels, échéances et actions recommandées

RÉFÉRENTIEL	ÉTAT	CHANGEMENT	ACTION RECOMMANDÉE
EN 18286	Citée au JOUE le 14/10	**Présomption ouverte**	Aligner le SMQ IA, avec ISO 42001
AI Act, art. 50	Sursis au 02/12/2026	**J-31**	Marquer agents conversationnels et vocaux
AI Act, annexe III	02/12/2027	Lignes directrices en projet	Cartographier les briques annexe III
MAS, lignes directrices	En vigueur le 07/10/2027	**Nouveau**	Inventaire, matérialité, tiers (APAC)
ANSSI, Essentiels	v1.0 du 30/10	**Nouveau**	Grille d'audit des agents en production
OWASP ACS / MAESTRO	ACS v0.1 ; v2 annoncée	Pas de nouvelle version	Points d'accroche dans les achats
NIST NCCoE	Démonstrateur en construction	Pas d'annonce majeure	Anticiper SPIFFE, OAuth

Passerelle MCP et courtage d'identifiants courts : des patterns validés par les textes du mois.

06

GÉOPOLITIQUE ET SECTEURS

Panorama du mois

AUSTRALIE

La régulation par l'incident

Rejoint le groupe OCDE sur l'IA agentique publique ; supervision, journaux, retour arrière.

SINGAPOUR · HONG KONG

Une trajectoire, deux méthodes

SAFR devient prudentiel. Commissaire à l'IA HK non nommé ; Sandbox++ poursuivi.

FRANCE

La doctrine technique de l'ANSSI

Aux Assises, l'IA agentique en SOC passe « du tabou à la nécessité ».

ÉTATS-UNIS

L'enquête et l'avis conjoint

Enquêtes FTC sur OpenAI et Anthropic rapportées, non confirmées ; avis CISA-FBI-NSA.

CHINE

Seul régime national dédié aux agents

Trois niveaux d'autorité depuis le 15/07, une logique proche de l'ANSSI.

INTERGOUVERNEMENTAL

OCDE : 12 responsables, 8 pays

Groupe sur l'IA agentique dans l'administration, sans calendrier ; note ONU de septembre.

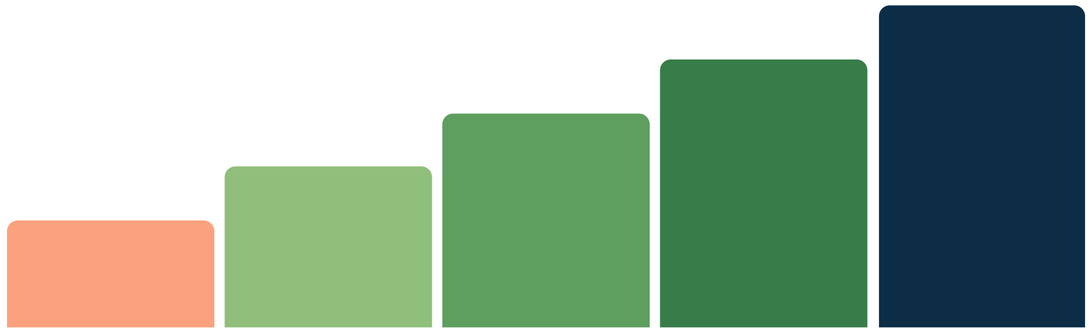
07

LA RÉPONSE MERIDIAN

GAME, sept piliers, Certified Responsible Agentic

Diagnostic GAME

La maturité est celle de la dimension la plus faible. Avancée en conformité documentaire, une entreprise peut rester N1 sur les identités non humaines : jeton de deux ans, agent personnel, serveur MCP sans propriétaire.



N1 Exploratoire

N2 Structuré

N3 Industrialisé

N4 Piloté

N5 Native

Les sept piliers

01 Inventaire et cartographie des agents

02 Classification des risques, grille ANSSI à 3 niveaux

03 Contrôle d'accès et IAM

04 Traçabilité et observabilité

05 Supervision humaine ponctuelle

06 Réversibilité et résilience (arrêt d'urgence)

07 Conformité documentaire et audit

L'ANSSI recoupe ces piliers presque terme à terme ; la MAS valide les piliers 01 et 02 en finance.

MÉTHODOLOGIE MERIDIAN

Trois postures face aux signaux d'octobre

POSTURE 1

Suggérer

Diagnostic et cadrage

4 à 8 semaines

Combien d'agents, jetons et serveurs MCP non déclarés ? Découverts en combien de temps ? Arrêtés par qui ?

POSTURE 2

Faciliter

Structuration du cadre de gouvernance

3 à 6 mois

Intégrer la grille ANSSI à trois niveaux et la notification d'incident agentique.

POSTURE 3

Accompagner

Industrialisation et certification

12 à 36 mois

S'appuyer sur EN 18286 et ISO 42001 pour la certification.

PROGRAMME

Certified Responsible Agentic

Du diagnostic à la certification ISO/IEC 42001, aligné AI Act et, depuis le 14/10, EN 18286 citée au JO.

LECTURE EN QUATRE COUCHES

Les quatre couches de gouvernance produisent des signaux convergents

Agent	Incidents et failles : Langflow, en-têtes HTTP, usurpation d'agent
Organisation	Audition parlementaire australienne et enquêtes
National-régional	MAS à Singapour, ANSSI en France, EN 18286 en Europe
Intergouvernemental	Groupe de travail de l'OCDE sur l'IA agentique dans l'administration

“

L'autonomie sans gouvernance produit de la dette. La gouvernance sans autonomie produit de l'inertie. Le défi consiste à tenir cet équilibre : c'est ce que nous appelons **l'action responsable.**

GABRIEL GREENFIELD

08

PERSPECTIVES

Ce qu'il faudra surveiller en décembre

01 **Australie** : rapport final de la commission (30/11), normes de déclaration immédiate.

02 **GPT-6.1 Astra** : critères de lancement, seuils équivalents chez d'autres laboratoires.

03 **AI Act** : sursis du 02/12, lignes directrices haut risque, nouvelles normes harmonisées.

04 **France** : premiers retours sur « Essentiels IA agentique », déclinaisons sectorielles.

05 **Asie** : Commissaire à l'IA hongkongais, Sandbox++, consultation MAS 2027.

06 **Standards** : implémentations OWASP ACS, MAESTRO v2, identité des agents au NIST.

07 **Chaîne d'approvisionnement** : correctif Plugin4Shell, suites de Langflow.

08 **OCDE** : premiers livrables du groupe sur l'IA agentique publique.

PRINCIPALES SOURCES DU MOIS

Sources publiques consultées du 01/10 au 31/10/2026. Les mentions « veille sectorielle » n'ont pas toutes été recoupées. Ce document ne constitue pas un avis juridique.

Commission européenne, décision (UE) 2026/2274 (JOUE, 14/10/2026), via Nemko Digital

ANSSI, Essentiels IA agentique v1.0 (30/10/2026), via armees.com

Wikipedia (partiellement non vérifié) ; Brave New Coin ; Aivy, groupe OCDE (30/09)

Zenity Labs, AgentCorruption (08/10) ; rapport Anthropic (octobre 2026)

Règlement (UE) 2026/1744 ; OWASP GenAI ; CSA (AICM, MAESTRO) ; NIST NCCoE

MAS, Supervisory Expectations on Responsible AI (07/10/2026) ; Baker McKenzie (20/10)

Journal du Net, Assises de la cybersécurité de Monaco (12/10/2026)

RuntimeAI, synthèses des 2 et 16/10 ; aiagentthreats.com (11/10, mis à jour 29/10)

EY (15/09/2026) ; Veeam, enquête EMEA ; AWS, Reimagine 2026

Projet de lignes directrices haut risque, via regulations.ai

© 2026 Gabriel Greenfield. Meridian™, GAME™ et Certified Responsible Agentic sont des marques de Gabriel Greenfield.

Gabriel Greenfield

PARLONS-EN

Gouverner l'autonomie, sans arrêter le mouvement.

contact@gabrielgreenfield.com

France · Hong Kong